

Zadanie 5 - Algorytmy genetyczne (optymalizacja)

Marcin Pietrzykowski

mpietrzykowski@wi.zut.edu.pl

wersja 1.0

1 Cel

Celem zadania jest zapoznanie się z Algorytmami Genetycznymi w celu rozwiązywanie zadania optymalizacji. Zadaniem, które należy rozwiązać będzie dyskretny problem plecakowy. Program jest do napisania w środowisku Matlab. Osoby chętne chcące napisać ww. program w innym języku odsyłam do punktu 5. Wszystkie informacje zostały podane na zajęciach. Przypomnienie najważniejszych informacji znajduje się poniżej.

2 Dyskretny problem plecakowy

Problem plecakowy często przedstawia się jako problem złodzieja rabującego sklep – znalazł on N towarów; i -ty przedmiot jest wart v_i oraz posiada określoną objętość c_j . Złodziej dąży do zabrania ze sobą jak najwartościowszego łupu, przy czym nie może przekroczyć objętości plecaka C_{max} . Nie może też zabierać ułamkowej części przedmiotów (byłoby to możliwe w ciągłym problemie plecakowym).

Formalna definicja problemu jest następująca. Do dyspozycji posiadamy I przedmiotów, a każdy z nich jest opisany jako para (v_i, c_i) :

$$A = \{(v_i, c_i)\}_{i=1\dots n} \quad (1)$$

Zadaniem jest znalezienie takiego podzbioru B , w którym wartość przedmiotów jest maksymalna, ale ich objętość nie przekracza maksymalnej objętości plecaka, czyli:

$$B \subset A, \quad \sum_{(v_i, c_i) \in B} v_i \longrightarrow \max, \quad \sum_{(v_i, c_i) \in B} c_i \leq C_{max} \quad (2)$$

3 Algorytm Genetyczny

Algorytm genetyczny jest to rodzaj heurystyki przeszukującej przestrzeń alternatywnych rozwiązań problemu w celu wyszukania rozwiązań najlepszych. Sposób działania algorytmów genetycznych przypomina zjawisko ewolucji biologicznej. Obecnie zalicza się go do grupy algorytmów ewolucyjnych.

Środowisko, w którym istnieje pewna populacja osobników, jest zdefiniowane przez problem, który należy rozwiązać. Każdy z osobników ma przypisany pewien zbiór chromosomów składający się z genów. Jest to genotyp danego osobnika i stanowi on podstawę do utworzenia fenotypu. Fenotyp to zbiór cech podlegających ocenie funkcji przystosowania modelującej środowisko. Innymi słowy - genotyp opisuje proponowane rozwiązanie problemu, a funkcja przystosowania ocenia, jak dobre jest to rozwiązanie. Każdy z osobników prezentuje lepsze lub gorsze rozwiązanie danego problemu. Każdy z osobników może posiadać kilka chromosomów, jednak bardzo często w algorytmach genetycznych wybrany osobnik posiada jeden chromosom. W literaturze oba pojęcia: chromosom oraz osobnik bardzo często używane są wymiennie.

Algorytm genetyczny posiada następujące kroki:

1. Wybór populacji początkowej
2. Ocena przystosowania osobników w populacji
3. Wyprowadzenie najlepszego osobnika

4. Sprawdzenie warunków zatrzymania
5. Selekcja chromosomów
6. Zastosowanie operatorów genetycznych
7. Utworzenie nowej populacji

3.1 Kodowanie informacji

Każdy chromosom koduje pewien zbiór informacji. Na potrzeby niniejszego zadania rozpatrzmy chromosom binarny, w którym każdy z genów może przyjmować tylko jedną wartość 0 lub 1.

Każdy chromosom będzie prezentowała jedno możliwe rozwiązanie tj. pewien podzbiór przedmiotów, które mogą zostać włożone do plecaka. Długość chromosomy będzie równa liczbie wszystkich przedmiotów I . Pojedynczy gen będzie definiował, czy dany przedmiot znajduje się w plecaku (wartość genu 1) czy też danego przedmiotu w tym plecaku nie ma (wartość genu 0). Przykładowy chromosom, dla zadania w którym występuje 15 przedmiotów, prezentuje diagram poniżej.

| | | | | | | | | | | | | | | | |
|---------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|
| nr przedmiotu | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| chromosom | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |

Chromosom koduje informacje, o tym że w plecaku znajdują się przedmioty: 1, 2, 6, 8, 14

3.2 Operatory genetyczne

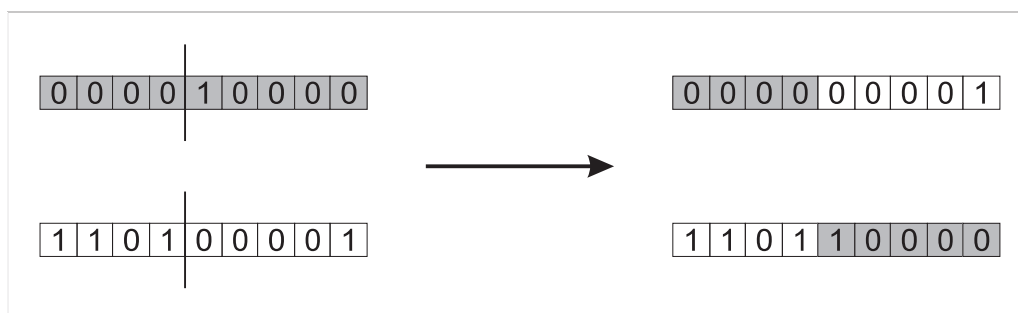
Operatory genetyczne mają na celu rekombinację genów w chromosomach. Wyróżniamy dwa operatory genetyczne: krzyżowanie, mutację. Niekiedy podaje się jeszcze trzeci operator - inwersja. Najbardziej popularne i najczęściej stosowane są dwa pierwsze. Każdy z operatorów genetyczny wykonywany jest z pewnym prawdopodobieństwem, które definiujemy na początku programu i które jest stałe w trakcie jego trwania. Typowe zakresy prawdopodobieństw dla poszczególnych operatorów genetycznych są następujące:

- Krzyżowanie 0.5 – 1.0
- Mutacja 0.0 – 0.1

3.2.1 Krzyżowanie

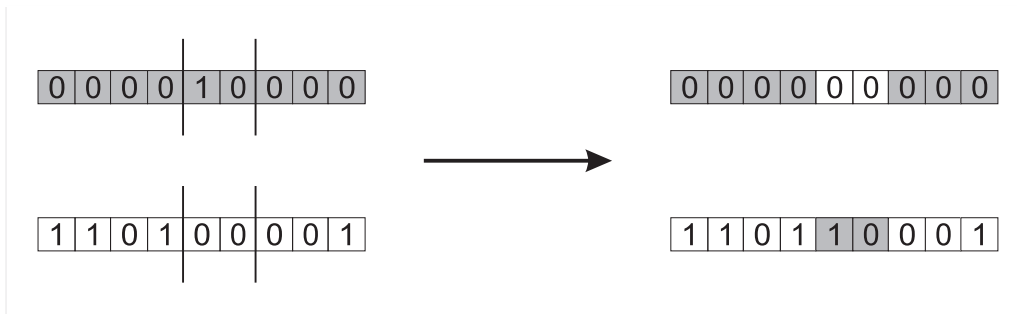
Krzyżowanie jest operacją mającą na celu wymianę materiału genetycznego między osobnikami.

Krzyżowanie jednopunktowe Krzyżowanie to polega na wylosowaniu jednego punktu krzyżowania, a następnie wymianie materiału genetycznego w następujący sposób:



Rysunek 1: Przykład krzyżowania jednopunktowego.

Krzyżowanie dwupunktowe Krzyżowanie dwupunktowe jest wykonywane analogicznie do krzyżowania jednopunktowego z tą różnicą, że losowane są dwa punkty:



Rysunek 2: Przykład krzyżowania dwupunktowego.

3.2.2 Mutacja

Mutacja ma na celu wprowadzić różnorodność genetyczną populacji. Prawdopodobieństwo mutacji wylicza się dla każdego genu. Mutacja w chromosomie binarnym polega na zmianie bitu na wartość przeciwną.



Rysunek 3: Przykład mutacji, która zaistniała dla 3 genów.

3.3 Funkcja przystosowania

Funkcja przystosowania ocenia jak dobrze dany osobnik jest przystosowany do środowiska (z punktu widzenia rozwiązywanego problemu). Dla dyskretnego problemu plecakowego będzie to wartość przedmiotów, jeżeli suma ich objętości nie przekracza maksymalnej objętości plecaka C_{max} , w przeciwnym razie będzie to 0. Matematycznie funkcję przystosowania dla danego osobnika można zapisać jako:

$$f(B) = \begin{cases} \sum_{(v_i, c_i) \in B} v_i, & \sum_{(v_i, c_i) \in B} c_i \leq C_{max} \\ 0, & \sum_{(v_i, c_i) \in B} c_i > C_{max} \end{cases} \quad (3)$$

3.4 Funkcja selekcji

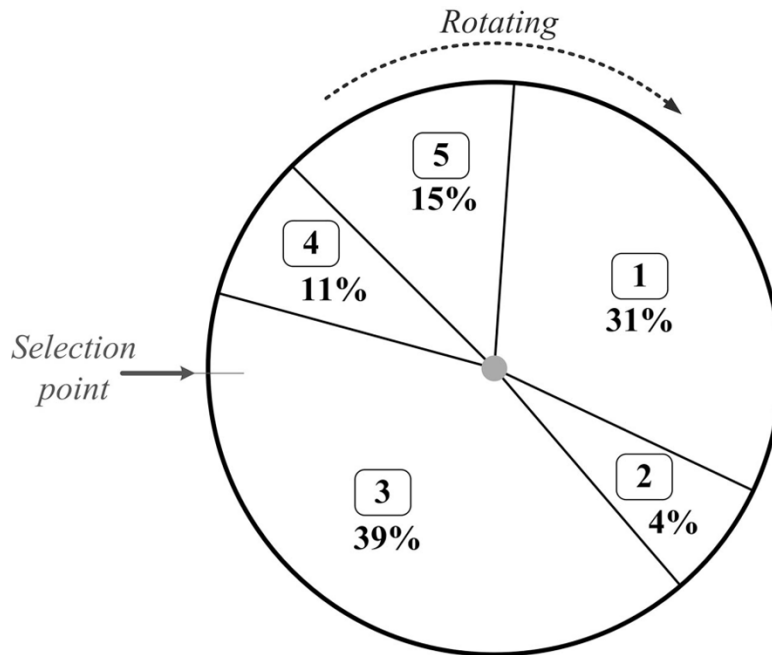
Metoda selekcji ma za zadanie wyselekcjonowanie najlepiej przystosowanych osobników przy jednoczesnym zachowaniu różnorodności genetycznej populacji. W wyniku selekcji dany osobnik może zostać wylosowany więcej niż jeden raz.

3.4.1 Selekcja koła ruletki

Polega na budowaniu wirtualnego koła, którego wycinki odpowiadają poszczególnym osobnikom. Im lepszy osobnik, tym większy wycinek koła zajmuje. Zajętość koła dla chromosomu ch_i w populacji złożonej z J osobników definiujemy z następującego wzoru:

$$p(ch_i) = \frac{f(ch_i)}{\sum_{j=1}^J f(ch_j)} \quad (4)$$

Rozmiar wycinków może zależeć od wartości funkcji oceny, jeśli wysoka wartość oceny oznacza wysokie przystosowanie. W takim układzie prawdopodobieństwo, że lepszy osobnik zostanie wybrany jako rodzic, jest większe. Niestety ewolucja przy takim algorytmie z każdym krokiem zwalnia. Jeżeli osobniki są podobne, to każdy dostaje równy wycinek koła fortuny i presja selekcyjna spada. Algorytm słabiej rozróżnia osobniki dobre od słabszych.



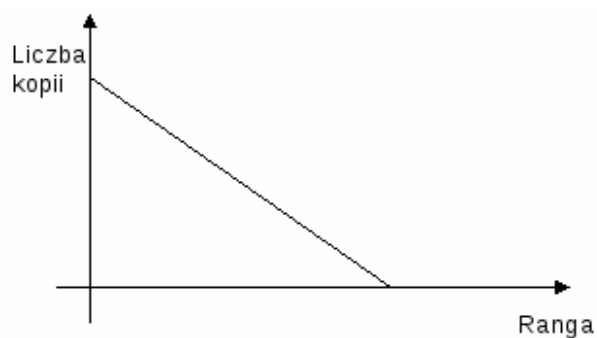
Rysunek 4: Przykład selekcji koła ruletki.

3.4.2 Selekcja turniejowa

Polega na wylosowaniu grupy k -elementowej z populacji a następnie wyborze osobnika o najlepszym przystosowaniu. Całą operację powtarzamy tyle razy ile jest osobników w populacji.

3.4.3 Selekcja rankingowa

Osobniki ustawiane są kolejno zgodnie z wartością funkcji przystosowania - od najlepszego do najgorszego. Każdy osobnik ma numer określający jego pozycję na liście, czyli swoją rangę. Liczba kopii każdego osobnika wprowadzana do populacji jest zdefiniowana przez wcześniej ustaloną funkcję, która zależy od rangi osobnika.



Rysunek 5: Przykład funkcji określającej zależność liczby osobników w nowej populacji od wartości rangi.

4 Polecenia

4.1 Na zajęciach

- Napisać skrypt do losowania problemu plecakowego dla zadanej liczby przedmiotów - n .

- Napisać ogólny skrypt realizujący ewolucję algorytmu genetycznego. Parametrami dla skryptu powinny być m.in.: rozmiar populacji, liczba iteracji, wskaźnik na funkcję przystosowania, wskaźnik na funkcję selekcji, wskaźnik na funkcję krzyżowania, wskaźnik na funkcję mutacji.
- Napisać funkcję (skrypt) obliczającą dla całej populacji wektor wartości przystosowania dla problemu plecakowego.
- Napisać funkcje realizujące: selekcję ruletkową, krzyżowanie jednopunktowe, mutację (wszystkie te funkcje powinny przyjmować na wejście całą populację).

4.2 Do domu

- Napisać dodatkowe dwie funkcje selekcyjne realizujące: selekcję rankingową i turniejową.
- Napisać funkcję do krzyżowania dwupunktowego.
- Dla wylosowanego problemu plecakowego przeprowadzić ewolucję algorytmu genetycznego. W każdej iteracji odnotować: (1) średnie przystosowanie populacji, (2) przystosowanie najlepszego osobnika w danym pokoleniu, (3) przystosowanie najlepszego osobnika wykrytego w dotychczasowej historii. Wielkości (2) i (3) nie zawsze muszą się pokrywać. Wartości te powinny zostać przedstawione na wykresie.
- Napisać skrypt rozwiązujący problem plecakowy poprzez indukcję (opartą na kracie) w sposób dokładny - zgodnie z wiadomościami podanymi na wykładzie. Dla wylosowanego problemu plecakowego (lub kilku) sprawdzić, czy algorytm genetyczny zwraca ten sam wynik, co rozwiązanie dokładne.

5 Wersją do osób nie lubiących Matlba

Osoby chętne mogą napisać ww. program w języku innym niż Matlab jednakże niezbędne będzie wyświetlenie wykresów. Oprócz zastosowania zewnętrznej biblioteki natywnej dla danego języka rysującej wykresy istnieje możliwość skorzystania z Matlaba jedynie w celu rysowania wykresów bezpośrednio w kodzie programu. Przykład jak można to wykonać w aplikacji konsolowej napisanej w języku C# jest do pobrania tutaj. Zaprezentowane rozwiązanie wymaga instancji Matlaba zainstalowanej na maszynie, na której uruchamiany jest program.